Review of Fishery-Independent Survey Programs In Southeastern U.S. Atlantic Waters

Beaufort Laboratory
National Oceanic and Atmospheric Administration (NOAA)
Southeast Fisheries Science Center
Beaufort, North Carolina

March 15, 2012

Abstract

A select panel of experts was assembled at the NOAA Beaufort Laboratory at Beaufort, North Carolina, February 28-March 2, 2012, to review state and federal systems for collecting fishery-independent data on reef fishes in the Exclusive Economic Zone of the South Atlantic bight offshore of North Carolina, South Carolina, Georgia, and Florida. The South Carolina Department of Natural Resources operates one system funded through MARMAP and SEAMAP programs. The Beaufort Laboratory operates another using NOAA funding. Historically three gear types have been used by South Carolina: trap (recently with video), short longline, and long longline—each gear is respectively used farther offshore. Since 2010 the Beaufort Laboratory has been using trap/video gear with methods identical to those used by South Carolina. The review panel primarily concerned itself with the effectiveness of gear types; procedures for randomizing sight and sample selections; spatial, temporal, and biological aspects of sampling; operational survey procedures; addition of sampling sites through reconnaissance; and data modeling and analysis. Variance components, computed within and between sampling events, and spatial autocorrelation of data across spatial scales, were explored as a basis for making recommendations. In depth statistical explorations, however, were beyond the scope of the review and therefore are included in the report as a prioritized list of research needs. In general, the panel recommended shifting effort and funding from long longline sampling to other gear usage, making short longline gear sampling biannual rather than annual, and extending its coverage in the study region to span the length of the continental shelf break. Trap/video sampling was recommended annually throughout the study area using spatial strata based on depth and latitude with the goal of homogeneity of fish abundance within a stratum. The establishment of a multi-agency steering committee was recommended that could convene annually to plan and coordinate between state and federal survey efforts.

Contents

Review of Fishery-Independent Survey Programs
In Southeastern U.S. Atlantic Waters

Introduction

A select panel of four experts met at the NOAA Beaufort Laboratory February 28 – March 2, 2012, to review fishery-independent sampling programs targeting reef fish in the U.S. South Atlantic. For fishery management purposes, the region is defined as federal waters extending three to 200 miles into the Atlantic Ocean off the coasts of North Carolina, South Carolina, Georgia and Florida. Fishery-independent sampling in this region is cooperatively conducted by state and federal entities. South Carolina Department of Natural Resources (SCDNR) conducts sampling through the efforts of the Marine Resources Monitoring, Assessment and Prediction (MARMAP) program of the Marine Resources Research Institute (MRRI). SCDNR also receives funding through the Southeast Area Monitoring and Assessment Program (SEAMAP)--a state/federal/university program for collection, management and dissemination of fishery-independent data in the southeastern United States. MARMAP began more than three decades ago and in recent years has extended its range of sampling to include federal waters from Cape Lookout, North Carolina, to Cape Canaveral, Florida. The Beaufort Laboratory of the Southeast Fisheries Science Center began fishery-independent sampling in 2010. Laboratory staff conducts sampling through a new effort named the Southeast Fishery Independent Survey (SEFIS) program.

MARMAP/SEAMAP and SEFIS collectively provide fishery-independent data for stock assessments conducted through the South East Data, Assessment, and Review (SEDAR) program of the South Atlantic Fishery Management Council. SEDAR was initiated in 2002 as a participatory process that involves scientists and data from multiple state and federal agencies, universities, and applicable non-governmental organizations. The goal is to improve the quality and reliability of fishery stock assessments in the South Atlantic through resource management based on sound scientific information. Traditionally stock assessments relied heavily on data from the commercial and recreational fisheries, but in recent years as fishery-management restrictions constrained both fisheries spatially and temporally, more emphasis has been directed towards fishery-independent surveys that use standardized research methodologies and are not constrained in time and space by regulations.

Aleta Hohn, Director of the SEFSC Beaufort Laboratory, addressed the select panel with an opening statement that emphasized the importance of this review. She stated that the outcome of the review is to produce statistical and logistical recommendations that could be instituted for the 2012 sampling season to improve the efficiency and effectiveness of the surveys. Todd Kellison, Beaufort Laboratory Fisheries Ecosystem Branch Chief, emphasized the goal of having a panel-produced report in hand by the close of business, March 2, 2012.

Panel

- John Walter, Research Scientist, NOAA SEFSC Miami Laboratory stock assessment biologist. He has a background in spatial analysis and recent experience with highly-migratory species assessments.
- Mary Christman, University of Florida, Adjunct Professor and Statistical Consultant. She holds a master's degree in marine biology and a PhD in statistics. She specializes in statistical sampling strategies, environmental and ecological statistical modeling, and the modeling of rare or clustered species.
- David Somerton, Research Scientist, NOAA Alaska Fisheries Science Center (AFSC), Seattle, Washington. He manages FSC at-sea trawl surveys in the Pacific Northwest and has worked on survey catchability and history in Hawaii for the management of the snapper-grouper and longline fisheries.
- Jeff Buckel, Associate Professor, North Carolina State University, Morehead City, North Carolina. He serves as a SAFMC Science and Statistical Committee member and has expertise in estimating vital rates and indices of abundance of fishes.

Larry Massey of the SEFSC in Miami served as panel chair. Warren Mitchell and Michael Burton served sequentially as rapporteurs. Also in attendance to provide expert information were:

- Marcel Reichert. He leads the MARMAP program for the South Carolina Department of Natural Resources and is a SAMFC Science and Statistical Committee member. In addition he was accompanied by program colleagues who provided additional information to the panel: Dave Wyanski, Survey Chief Scientist and head of The Reproductive Laboratory; Lori DiJoy with expertise in life history and sampling design; Michelle Pate with expertise in data management; Tracy Smart, Survey Chief Scientist with expertise in assessment data; and Dawn Glasgow, resident PhD candidate with expertise in video analysis methods. These represent some of MARMAP's 13 full-time state employees, 3-5 seasonal staff, and resident students from the College of Charleston.
- Todd Kellison, Beaufort Laboratory Fisheries Ecosystem Branch Chief in charge of fishery-independent surveys.
- Nate Bacheler, SEFIS Research Coordinator and Survey Chief Scientist. In addition, the following SEFIS researchers were on hand to provide expert information: David Berrane with expertise in sampling techniques and logistics; Neil McNeil with expertise in fishery video reading techniques; Warren Mitchell, Survey Chief Scientist with expertise in sonar, mapping, and fishery techniques; Christina Schobernd with expertise in data management, GIS, and fishery video; Zeb Schobernd, Survey Chief Scientist with expertise in fishery video sampling; and Lisa Wood with expertise in GIS and spatial data analysis. SEFIS includes 1 federal full-time employee, 4 full-time contractor staff, and two full-time-equivalent staff working part-time on SEFIS objectives.
- Erik Williams, Beaufort Laboratory Sustainable Fisheries Branch Chief. He leads the SEFSC fishery stock assessment team for the South Atlantic stocks managed by the SAFMC.
- Kyle Shertzer, Research Scientist, Beaufort Laboratory, with extensive experience in South Atlantic stock assessments.

Presentations

To brief the panel on all aspects of fishery-independent sampling in the U.S. South Atlantic, the following presentations were made the first day of the meeting with extensive questions and answers between panelists and presenters:

- Marcel Reichert, "The MARMAP/SEAMAP Reef Fish Program: An Overview."
- Todd Kellison, "SouthEast Fishery-Independent Survey."
- Marcel Reichert, "MARMAP Sampling Gear and Deployment."
- Marcel Reichert, "Shipboard [fish sample] processing" and "Laboratory processing and analysis."
- Nate Bacheler, "Use of underwater video to index reef fish species by MARMAP/SEAMAP/SEFIS."
- Nate Bacheler, "Standardizing reef fish trap CPUE or video counts using delta-GLM models."
- Erik Williams, "Abundance data for stock assessment use in the US South Atlantic."
- Christina Schobernd, "South Atlantic Habitat Information: Knowledge and Gaps."

Definitions

To maintain clarity in discussing statistical aspects of the review, the panel provided the following definitions:

- *Study area:* The area extending three to 200 nm into the Atlantic Ocean off the coasts of North Carolina, South Carolina, Georgia and Florida.
- *Sampling universe:* The population over which the survey desires to make inference. This is the entire hard bottom habitat in the *Study area*. Unfortunately the entire sampling universe is unknown at present and will be for the foreseeable future.
- *Sampling frame*: The universe of sites from which sample sites are drawn, i.e., point locations of known hardbottom.
- *Site*: The primary unit of sample selection. This is one individual location from the sampling frame.
- *Random sample*: The initial set of *N* sites chosen from the sampling frame.
- *Recon sites*: Sites which were not in the original year's *random sample* but that can be opportunistically sampled for the purposes of identifying new sites to be added to the *sampling frame*.
- *Opportunistic sample*: A site not part of the original random sample but part of the sampling frame that was sampled in a given year.
- *Event*: A group of *sites* within close proximity. Usually the initial random site plus the 5 additional sites from the *random sample* selected prior to the survey.
- *'Non-response' bias*: Bias imparted by not indexing areas not in the sampling frame, e.g. artificial reefs, areas north of Cape Lookout, and unknown hard bottom.
- *Stratified sampling*: A stratification scheme whereby the study area is partitioned into strata that are as homogenous as possible with respect to fish abundances; samples are collected in every stratum in every year.
- *Grid cell:* A single unit (stratum) of the stratification scheme.
- *Sampling design*: The method of sample selection, often probability-based (e.g. random, stratified random) for *design-based* estimation, where the means and variances are functions strictly of the sampling design. For *model-based* estimation where means and variances are functions of the model, strict probability-based sampling potentially can be relaxed. The key element is that the set of samples are representative for the *sampling universe.*
- *Critical uncertainty areas:* Grid cells that have a high probability of hard bottom, based upon a suite of GIS layers and probability mapping, but for which there is little survey coverage.
- *Relative standard error (RSE)*: Standard error of the estimate divided by the estimate. Note that this is not the coefficient of variation which is defined as the population standard deviation divided by the population mean.
- *Sample size sensitivity study:* A recommended study of the sensitivity of the model-based RSEs and the RSEs for age composition metrics to the absolute sample size. This will inform policy makers of the trade-offs between reduced funding and increased scientific uncertainty.

- *Standardization models*: statistical models for standardization of catch or catch rate data. These models are essential to interpretation of CPUE trends and variances under a model-based sampling approach. Their primary practical purpose is to adjust for an unbalanced sampling design (unbalanced with respect to the explanatory variables), account for unaccounted factors in the sampling process (e.g. temperature) and properly account for variance and covariance structure in the data.

The Review

The stated objective of the review was "to provide (1) an independent review and assessment of, and (2) recommendations for, the function and integration of the fishery-independent survey programs targeting reef fish in South Atlantic waters. Specifically, we are seeking guidance on how to maximize the efficiency and usefulness of information generated by current surveys. Recommended changes to current survey methodology and approaches should be feasible given current multi-program funding and logistical constraints, and should balance the benefits of potential changes to survey methodology with anticipated impacts of interrupting existing long-term survey time series."

Two sampling systems were under review. One is state operated; the other is federally operated:

- The South Carolina Department of Natural Resources Marine Resources Monitoring, Assessment and Prediction (MARMAP) program / Southeast Area Monitoring and Assessment Program (SEAMAP) Reef Fish Survey (hereafter MARMAP/SEAMAP).
- The NMFS-SEFSC South East Fishery-Independent Survey (SEFIS).

The panel review was provided terms of reference or issues to be considered. It consisted of six categories, each category containing specific questions or statements. Each is addressed below:

A. Survey Utility:

To what extent are data generated from MARMAP/SEAMAP (trap, video, still pictures, short longline, long longline, and hook and line) and SEFIS (trap & video) surveys utilized, or likely to be utilized, in stock assessments or to address other management needs? How could the utility of surveys be improved?

Short longline—Data from the short longline survey are currently not used in any current stock assessment, but have potential for such use for snowy grouper and speckled hind, which inhabit the targeted shelf edge / ledge habitat. The primary shortcoming of the survey is that its spatial distribution, between 32°N and 34°N, is likely inadequate to cover the spatial distribution of key species, particularly snowy grouper and red snapper. If sufficient resources cannot be obtained to expand the latitudinal range of the current survey, we recommend pooling resources over time and conducting more spatially comprehensive surveys in alternate years. If such a spatial expansion would require the use of another vessel, in addition to the RV Palmetto, the review committee is concerned that the skill level of the new vessel to fish in this difficult habitat may not be sufficient to avoid a vessel effect in the data unless considerable effort is expended in standardizing fishing techniques.

Long longline—Data from the long longline survey, which targets tilefish, were used for tilefish assessment, but the catch rates are so low that their information content was deemed substantially lower than the fishery-dependent index. Consequently this survey is not likely to play an important assessment role. One reason for this is that the survey was initiated as an exploratory fishing operation to determine if the stocks of tilefish could support a northward extension of the commercial fishery, so that the survey spatial distribution is disjointed from the bulk of the fishery. The initially low catch rates remain low and the commercial tilefish fishery remains in Florida. However, if it was considered necessary to continue a survey for this species, some type of industry partnership should be considered. This could be funded by a research set-aside of some fraction of the quota similar to what is done for Atlantic sea scallops, Pacific sablefish and other species. In addition, it may also be possible to obtain funds from the NMFS fisheries Cooperative Research Program to help fund the survey, similar to the monkfish trawl survey conducted by the NEFSC. This approach shares the costs of assessment between the fishery and the scientific agencies.

Hook and line—Survey personnel should be commended for the proactive efforts to collect and process data that will facilitate ecosystem management. The data from the hook and line collections are primarily for biological material, diet and life history studies, but currently plays a very small role in the stock assessment process.

Trap survey—Data from the MARMAP trap survey are currently used in several fish stock assessments; however there are several shortcomings of the survey that limit the utility of the data. The most important of these is the spatial coverage of the survey. The survey index of abundance is based on the premise that the population trends in the observed areas accurately reflect trends in the unobserved areas. The survey targets species associated with hard, live bottom habitat distributed in widely dispersed patches whose locations are incompletely known, but has not extensively covered the northern and southern extremes of the south Atlantic bight. The spatial coverage has been greatly improved by the addition of the SEFIS sampling in the southern part of the area, but there still remain significant under-sampled areas where commercial and sport catch and fisher knowledge indicates there may be habitat, especially in the far north. We believe that it is critical to expand exploratory operations to currently under-sampled areas with the objective of finding new areas of appropriate habitat and achieving a more representative spatial distribution of the trap sampling effort. This need is so important that some redirection of vessel time currently used for trap sampling might be better spent surveying for new sampling locations. This tradeoff is examined below.

Coupled with the likely spatial variation in the proportion of the habitat that is actually sampled, an additional shortcoming is the potential for uneven spatial coverage due to interruption of the survey by weather or other events. Although the sample allocation at the beginning of the survey is based on random sampling over the known distribution of live bottom habitat, the number of samples actually collected may be quite less than the initial allocation, leaving holes in the spatial pattern of final samples. We recommend a sampling strategy below that should help to alleviate some of these issues, while maintaining the integrity of the index.

We believe that it is imperative that a clearly written sampling manual be created, consistent with the NMFS Fixed Gear Survey Protocols Manual (NOAA 2003), that details trap specifications, bait, deployment procedures, site selection and all issues that could potentially affect trap catchability. This is to ensure consistency between the MARMAP and SEFIS data as well as the continuity of the data over time. Creation of such a document was once mandated by the Director of NOAA for all NMFS surveys and is standard for NMFS trawl and acoustic surveys.

Video surveys—Data from the video and still cameras attached to the fish traps are currently not used in any stock assessment and we feel that the video should not replace the trap without adequately addressing potential sources of bias and calibration of the two gears. In addition the sheer volume of processing time will make its use as an index limited in the short term, until more efficient reading technology can be incorporated. However this data could provide ancillary information to the stock assessment models that may lead to better predictions. Video, in conjunction with trap catches, may help to address biases due to species and size selectivity, saturation and incomplete detectability in trap catches. Video observations are also subject to biases due to the inability to directly measure and speciate the observed fish and environmental variability affecting viewing conditions. The issue with fish measurement can be partially addressed by using stereo video cameras; variability in viewing conditions can be addressed by measuring light level and water transparency or restricting counts to be within a specified distance from the camera. But with currently available technology the limiting factor to the utility of video data is the huge amount of time needed to view the videos and extract the data. The cost of video processing is repeatedly reported as a limiting factor at all NMFS labs that use video to obtain fish density estimates, although labs processing the videos only for the relative abundance by species or fish length are achieving greater success. The video data now being collected, however, can play a very important role in stock assessment models. Stock assessment models currently estimate selectivity, which is generally considered a function of the availability of the fish to the sampling gear and the size selectivity of the gear. If the size distribution of fish determined from the video for a single trap can be considered as representing the size distribution of fish available to the trap, then the size selectivity of the trap can be determined using models widely available for trawl and gillnet mesh selectivity (Millar 1992, Wileman et al. 1996). The empirical estimation of size-based selectivity could provide information on the functional form of selectivity and inform priors in Bayesian stock assessment models. Using informative Bayesian priors to constrain the values of selectivity parameters has been increasingly shown to produce better behaved model fits, often with more precise model outputs.

B. Survey design

General recommendations regarding survey design.

1) Currently the major assumption of these surveys is that the population trends in the observed areas accurately reflect trends in the unobserved areas, e.g. above Cape Lookout or on artificial reefs. This is a strong assumption that is critical to the acceptance of the survey by stakeholders. Below, we refer to the bias imparted by not indexing areas not in the sampling frame as 'non-response' bias. Related to this, we recommend:

    i. A spatial stratification scheme based upon latitude and depth for the existing sampling frame be developed. Further, we recommend that the stratification extend beyond the current latitudinal limits of the sampling frame to include the entire putative *sampling universe (*see *Critical uncertainty areas* for how to expand the putative *sampling universe*). Stratification should be based on relatively homogenous latitudinal partitions. A second level of stratification that incorporates depth within a latitudinal stratum should be included because abundances and species vary across depth gradients. For this reason the stratum may need to become narrower at the shelf edge and extend to the limit of the trap/camera sampling depths.

    ii. That the area occupied by known artificial reefs, wrecks and obstructions be quantified so that the area of the sampling frame can be compared against the total area of artificial reefs. This would allow one to make the statement that the current survey indexes *X* square kilometers whereas artificial reefs only occupy *Y* square kilometers. If these artificial structures are or become in the future a substantial proportion of total hard bottom area, perhaps they need to be considered for some type of sampling.

2) Based on a cursory assessment of the data for black seabass, the between-event variability is greater than the between-year variability, indicating that events on different reef complexes have greater variability than the interannual variability that we wish to capture. For this reason and because the end products are all model-based relative indices and relative age composition, we recommend that a set of fixed events be chosen to be a 'core' set of sentinel events that are

repeatedly sampled in every year. This would minimize the spatial variance which appears to dilute inter-annual signals.

3) This core set of fixed events would be determined by several criteria:

   a) Sufficient sample size such that the integrity of the current survey would not be compromised both in spatial representativeness and in RSE. The minimum sample size would be determined subject to *Sample size sensitivity study*.

   b) Identification can be at least partially determined by examining sampling areas that have routinely always received effort.

   c) The core events would need to cover the stratified sampling universe, i.e. at least one or more events should be in every stratum.

   d) These core events would be sampled as close as possible to the same latitude and longitude each year, maximizing detection of trend. If sampling cannot be completed at the same season each year than some type of seasonal stratification may be necessary or season may need to be considered in model-based CPUE estimation.

   e) In any given year, additional samples from the original sampling frame should be allocated proportionally to fill in spatial gaps; the extent to which a 'gap' exists may be informed by spatial autocorrelation.

4) These fixed locations would constitute the 'core' set of sentinel events to be sampled repeatedly year after year, augmented as funding and time allows with additional events. In fact the survey appears to have operated this way *de facto* anyway, but without the explicit statement of this design as many of the same events appear to get drawn in the random selection year after year.

5) Priority should be given to adding additional recon (and then, as these initial recon sites meet inclusion criteria, they should be added to the 'core' or sampling frame sites) sites in the North or to areas identified as those of *critical uncertainty*. Secondarily it may be necessary to add sites to the core in the North and South to ensure adequate spatial coverage.

6) Some considerations of the proposed fixed site strategy are as follows:

   a. Cons regarding fixed site:

      i. Ephemeral nature of some reefs.

      ii. Depletion over time could be a factor for some animals.

      iii. Does not capture population age composition necessarily.

   b. Pros regarding fixed site:

      i. Captures trend, minimizes spatial variance.

      ii. Greatest power to detect change, which is the key product to the assessments.

      iii. Generally will have a fixed minimum cost.

      iv. Captures relative age composition with improved precision.

7) It is imperative that there is clearly written sampling strategy that documents the exact sampling design from initial sample site selection to on site gear deployment within a selected sampling

site, selection of sites within an event as well as recon sites. This needs to be a recipe that could be followed and executed by any chief scientist. This is a formal NMFS requirement.

8) Subjective decisions about sample placement should be avoided and all decisions made objective. For instance when using multibeam mapping or the split beam echosounder to make decisions on whether recon sites should be made in a given area, there should be some rule that can be followed. This could be to take the first sign of hard bottom/reef habitat or that the habitat must meet some clearly definable criteria. If it is simply the decision of the chief scientist and the captain, there is no way to replicate this process or to protect from bias which might occur when there are two identified areas of hard bottom from which to sample and the decision gets made to go to the area of highest relief, most fish observed on the echosounder, etc.

Are the survey designs (for trap-video and longline surveys) used by MARMAP/SEAMAP and SEFIS most likely to minimize variance and bias across survey response variables, relative to other survey designs?

We feel that the trap/video design can be improved by stratification which would ensure spatial representativeness. The current sampling design may oversample in some areas; this can be determined by the *Sample size sensitivity study,* see below. Oversampling can occur when there is an overabundance of samples relative to the target index RSE or when many samples are in close proximity so that these samples are non-independent. If oversampling is present, then the extra samples could be reallocated to recon in the critical uncertainty areas.

The currently extremely high level of spatial variability poses two problems related to variance. The first is that variability between event locations appears to be much higher than interannual variability obscuring trend detection. Fixed sites may help remove some of the between-event variability. Second, the standardization models do not currently account for the fact that adjacent samples may be more alike and hence do not adequately account for covariance.

With regards to bias, the primary problem, as noted above, is the assumption of no 'non-response' bias. There is also the possible problem that current approaches for addition of new sample sites by recon may not protect from bias, although there are no signs of bias based on a very cursory observation. We recommend that this issue be further explored. One way to look for potential bias would be to examine stratum level means rather than overall means and this remains a topic that needs to be addressed as sites are added.

Should a cell-based sampling approach be considered? If yes, how should cell size be determined?

Yes stratification should be used to ensure representative allocation of samples in any given year as well as ensure complete spatial coverage. The first step would be to ensure at least one sentinel event in each stratum. As funding allows, random sampling sites should be added with proportional allocation based on the *Sampling frame*. Grid cell (stratum) size should be based upon the distance over which one can create relatively homogenous areas with respect to the fisheries of interest, these might be some latitudinal distance and some type of depth bin as described above. Final choice of strata should rest on priority species or some composite measure of likely community structure. Similar types of cells are used in the Gulf of Mexico surveys.

Considering potential spatial autocorrelation, should the experimental unit be a trap site or a group of trap sites?

We feel that the experimental unit should remain a *site* (individual trap) but that the correlation structure between traps within an event should be accounted for in the standardization model. From a variance components model of some species observed in the traps, it appears that the within event (a set of 6 or fewer traps) correlation is high relative to the between event correlation. This indicates that traps in close proximity to each other are not independent and represent pseudo-replicates. On the other hand, this correlation structure can be used to guide the appropriate sample spacing, that is, it may help in the spatial allocation of traps.

Can sites not selected randomly for sampling in a given year be sampled and included in an index of abundance? Currently non-randomly selected sites are included in an index.

We think that the new areas might be accounted for by appropriate statistical models. We feel that the *opportunistic* samples can be added as they are part of the original sampling frame if they were not subjectively chosen. Opportunistic sites should be chosen as haphazardly as possible. We feel that recon samples should not be included in the index.

The elevation of recon sites to the sampling frame makes a strong assumption that the trend in the unsampled areas is the same as the trend in the sampled areas; further it assumes that the effect of adding the new areas relative to the current areas can be adequately estimated with a few years of data. For instance this would assume that any latitudinal cline in abundance could be estimated in the standardization model if sampling were expanded above Cape Lookout. This assumption should be tested by examining year*area interactions in the models and the study of potential biases in recon sites, below.

How much effort should be utilized to identify new areas for sampling (e.g., mapping and obtaining information from commercial or recreational fishers)? How should such "new" areas be incorporated into the survey?

Spatial representativeness of the sampling universe is of paramount importance which will require additional recon sampling near and north of Cape Lookout. Nonetheless, maintaining the integrity of the current sampling and RSEs around 0.2 for key species restricts the amount of additional sample sites that could be reallocated to areas of critical uncertainty. The amount of effort will largely reflect funding but one means to determine how much effort can be given under current funding is to determine how many samples the current survey is oversampling. The number of samples that can be removed and still maintain a target RSE of 0.2 would then (in a level funding climate) free resources for recon. Allocation of recon samples should be done by identifying the grid cells that have the lowest level of current sampling and the highest probability of having hard bottom. Spatially explicit catch rate and effort data obtained from headboat and commercial observer data, recreational and commercial fishing knowledge and landings, Dunn and Halpin (2009), sediment probability maps, sediment data and multibeam mapping results can all be combined to develop a more refined hard bottom probability map.

How should such "new" areas be incorporated into the survey?

The area around and north of Cape Lookout should be included in the sampling frame as soon as possible through the detailed protocol of moving a site from recon into the frame.

How will these "new" areas be treated in terms of developing indices of abundance?

These 'new' areas will need to be dealt with in the statistical modeling. In general, as new areas get added, this will make nominal means inappropriate for trend analysis. These new areas will have to be new spatial cells in the stratification and this will require the assumption that the historical trends for the unsampled areas matches the sampled areas. (Note: for short bottom longline gear and long bottom longline gear, please see considerations specified above in the Survey Utility section.)

C. Statistical Treatment of the Data

Should the survey data be 'standardized'?

In order to obtain estimated mean CPUE for input to stock assessments or other information needs, we believe that standardization is necessary to ensure that annual CPUE estimates are based on the same sets of conditions. As new areas get added to the sampling universe within which data are collected, nominal means will become more and more inappropriate, especially for trend analysis. But nominal <u>stratified</u> means could be useful for cursory explorations.

If yes, what statistical models should be used?

Before discussing potential models, we have two important recommendations:

First, we recommend that the response variable of the statistical models be the observed counts in the traps, not the CPUEs (count/soaktime). The effects of differential soak time can more appropriately be accounted for in the standardization model by the use of an offset (cf. Cameron and Trivedi, 1998) that is included on the right hand side of the model, rather than as the denominator in the response. There are several statistical reasons for doing this including obtaining unbiased estimators with lower variance than when CPUE is used as the response (Cochran, 1977). Statistical modeling of the observed counts and using an offset for soak time when standardizing ensures that the adjusted means are the best linear unbiased estimators of mean CPUE. Conversely, use of CPUE rather than the counts increases the likelihood that the resulting adjusted mean is biased where the size of the bias is a function of the sample size.

Second, when using CPUE as the response, the underlying assumption is that the relationship between the observed count and soak time is linear. This is equivalent to claiming that the likelihood of an animal entering the trap does not change over time. This is not likely to be entirely true for reef fish as they interact and hence demonstrate non-constant rates of collection in traps. As a result, it is likely that the relationship is non-linear. The most obvious cause of non-linearities in the relationship could be saturation or species interactions. We recommend that the relationship between soak time and counts be further studied to identify whether there is some range of soak time over which the relationship is linear. Or, if not, a study might determine the best relationship to be used in the models. One approach would be to mine the current dataset for samples (individuals trap samples) that have similar characteristics, such as depth or season or geographical location, and, if possible, abundance for the species of interest, but which have variable soak times over the likely range that would occur in the field. The results of such a study could be used both in modeling the current data as well as informing future field work with respect to optimal soak times. If saturation does exist, statistical models for censored data may be able to deal with the downward bias in CPUE.

On a related note, we wish to point out that the probability of presence in a trap is also likely to depend on soak time, especially for the rare or more “shy” species. Hence, should the best modeling approach be determined to be a zero-inflated model (see below) then the Bernoulli portion of the model should also include an effect due to soak time.

This leads to the additional point that a critical aspect of maximizing use of the collected data is the need to identify the appropriate probability distribution for the response variable for use in the statistical modeling for standardization. The current approach of assuming a delta-lognormal mixture for the response variable is limiting in that it assumes that all observed zeros are structural zeroes. It has been shown that this is not the case for several species reviewed by Bacheler et al. (in review). They compared counts from video cameras attached to traps to the trap counts and found that for some species, the traps recorded zero individuals even though they were observed on the video. The delta-lognormal further relies on the above-mentioned assumptions that the observed counts are linear in soak time since it uses CPUE as the response variable. So, alternative distributions that handle the zeroes more appropriately should be explored, including count distributions (e.g. Poisson, Negative Binomial, quasi-Poisson), the zero-inflated (ZI) versions of these distributions (ZI-Poisson, ZI-Negative Binomial, ZI-quasi-Poisson), and other distributions that are sufficiently flexible to include large numbers of zeros such as the Tweedie distribution. The zero-inflated distributions, especially, may be of use in identifying the field conditions under which it is likely that a species should have been caught in the trap yet no individuals were caught. Combining this with further analyses of occupancy and detectability such as is being done currently by SEFSC personnel using Royle’s N-mixture approach (Royle, 2004) should inform the ability of the observed count from a trap or video to adequately represent the population (i.e. local abundance) being sampled by that trap or video. This is especially useful if the sentinel event approach is implemented.

In addition to identifying the best distribution for the observed counts from the trap (and possibly video) gear, we recommend that models include random and repeated measures effects to include several likely sources of correlations in the observations. These sources include temporal (observations taken from the same location over time are likely to be more similar to each other than to observations taken at different locations), spatial (synoptic observations nearer in space are likely to be more correlated than those farther apart in space), and event (traps within the same event set of 6 are likely to be correlated due to local environmental and habitat homogeneity). We recommend that the potential sources of correlation and the form of that correlation be explored before being included in the models. An important reason for including these effects in the model is that a model that ignores the correlation is underestimating the true variances and standard error of the means.

Even though we recommend incorporation of the random effects, we wish to make several cautionary comments. First, using random effects with non-normal distributions, such as the Poisson, may lead to model convergence problems. Second, the choice of probability distribution of the random effect must be carefully considered. The default in most statistical software is that the random effect is a Normal variate with mean 0 and constant variance. Instead there are alternative distributions; such models are called hierarchical models that could be explored. This is not meant to discourage exploration of inclusion of the random effects, but instead to ensure that the modelers correctly fit and interpret the data and results. Note that standardizing the CPUE from the sentinel sites alone would only require a repeated measures model to correctly incorporate time effects.

What factors should be included in the standardization (temperature, water clarity, depth, season, latitude, etc.)?

Because 1) the sampling frame is being expanded each year through recon sites, 2) the sampling each year did not necessarily cover the same range of influential covariate values, and 3) in light of the recommended emphasis on doing recon in geographic regions currently under-sampled, we recommend that CPUE be standardized. The following factors should be considered but may not necessarily be ultimately used: temperature, water clarity, depth, season, latitude, time of day (note that sunrise and sunset is later and earlier at depth than at the surface), current, and tidal cycle. This list is not necessarily exhaustive but represents conditions that we believe are likely to affect catchability or be unbalanced through time (i.e. for inter-annual comparisons). A variable need not be included under the stipulation that it always takes the same values and frequencies of those values every year. It is unlikely that these can be controlled from year to year and so any variable, even if not listed, should be considered for possible inclusion if it is known to have an effect on CPUE or if it cannot be controlled and so would vary significantly between years. Note that a variable that may drive abundance should never be used in the CPUE standardization because it will remove the effect of this factor on the annual signal. A case in point would be to not include a large-scale climatic factor such as the NAO, if it is thought that this might actually drive abundance from one year to the next.

Before these variables are included in the models, we suggest that the form of the relationship be completely explored so that incorporation is appropriately done. It is imperative to plot the relationship of the factor versus the dependent variable to explore the functional form (linear, threshold, non-linear, etc) and to inform how to model the factor, i.e. continuous or categorical. Classification and regression trees (CART) may be useful for identifying categorical break points. Generalized additive models (GAMs) can be used to inform possible non-linear relationships.

Related to their use in CPUE standardization, it may be necessary to *a priori* impose some functional form for an explanatory factor on the data even if it is not significant in the current model. This could be done if prior studies exist on the nature of the response variable versus the factor. Temperature, in particular should be explored via threshold models or some other non-linear form.

D. Survey gears

Are the gears and gear designs used by MARMAP/SEAMAP and SEFIS appropriate given survey objectives, constraints, and history (i.e., existing time series)?

We recommend maintaining all currently used fishing gear (not videos) in their current design and method of deployment, but believe that further research is needed to better understand how each gear type samples the available fish. Specific examples include: 1) longlines and traps can be subject to saturation when there is localized high fish abundance, which causes the relation between CPUE and true density to vary non-linearly. Catch rates from both gear types should be examined to determine whether saturation is occurring. If saturation is present, other definitions of CPUE that are less subject to saturation should be considered, 2) traps can have domed shaped selectivity indicating that larger fish are progressively under-sampled with increasing size. When this occurs, length frequency samples are less informative to stock assessment models, especially for estimating fishing mortality rate. However, this problem could be lessened if the shape of the selectivity function were known *a priori* or could be estimated from additional data. Video data might provide this needed additional data. We recommend the incorporation of new, or modifications to current, video gear that allows for measurements of the sizes of fish available to each trap, which if combined with the sizes of captured fish, should allow the estimation of a trap selection function for each species. 3) Video observations, especially those taken using a bait to attract fish, can provide a biased view of local density, species composition and even length composition. Part of this problem is related to the decreasing ability of correctly identifying and measuring fish with increasing distance. Consequently, it is important to develop estimators of CPUE that are robust to changes in light level and water clarity. Further research toward developing such estimators is very important while the video time series is still short.

E. Sampling methodology

How can subjectivity be minimized in terms of determining exactly where / when to deploy traps?

The current sampling methodology allows for trap sampling decisions to be subject to differences in cruise personnel. Potential biases include how the decisions of which site are chosen for initial selection, which additional sites comprise the event as well as the exact location of gear deployment. The former decision violates the random sampling protocol because it does not protect from non-response bias, i.e. the initial site decision may be made by weather, proximity or some ad hoc desire to ensure spatial representativeness.

There may be other subjective decisions that we have not listed here and we recommend that former and current chief scientists work together to populate a list of decisions (that they make during a cruise) and determine where subjectivity may play a role. All possible subjective decisions need to be objectified. For example, the decision on where to set the gear within a chosen location needs to be made entirely objective. If the latitude/longitude point does not have hard bottom then an explicit protocol should be followed to identify the spot where the trap is deployed (e.g., first hard bottom seen during a consistent search pattern, and determination of hard bottom needs to made objective, e.g. return strength). The presence of fish alone (i.e., without hard bottom) should never be used to determine trap deployment. This sampling methodology and ALL decisions need to be written as explicitly as possible to ensure sampling is replicable across cruise personnel and vessel.

We recognize that ultimately some subjective decisions will be made by the captain and chief scientist due to exigencies of the nature of the work. All decision points need to be clearly documented, e.g. the captain chose to set gear up current of reef, when the objective protocol cannot be followed.

We strongly recommend that a NOAA technical report (as outlined in the document: NOAA Fisheries Protocols for Fixed Gear Trap, Pot, and Camera Pod Surveys) be written that contains this explicit protocol so that differences in personnel do not influence index integrity

F. Program Coordination

How should survey effort be divided between MARMAP/SEAMAP and SEFIS?

The survey efforts of MARMAP/SEAMAP and SEFIS should be made by a steering committee. The steering committee should be composed of end users of the data including SEFSC stock assessment scientist(s) (e.g. Erik Williams and staff), SEDAR staff (e.g., Kari Fenske), SERO staff, a SAFMC member, and personnel responsible for running MARMAP/SEAMAP/SEFIS. The steering committee should do year-to-year coordination of surveys including site selection, training, and data collection and management as well as provide recommendations on responses to decreases in funding and allocation of resources. This committee would further ensure that lines of communication would be maintained between these different entities. Recommended changes to monitoring programs would be communicated to parties responsible for overseeing the leaders of MARMAP/SEAMAP (NOAA technical monitor) and SEFIS to ensure that changes are implemented without concern for modification of former agreed-upon objectives.

How can efficiencies be gained (e.g., shared versus separate data management efforts)?

In response to limited funding or funding reductions, the MARMAP/SEAMAP/SEFIS program needs to communicate the costs of reducing survey resources to users of reef fishery resources. When used in a probability based stock assessment, increased scientific uncertainty in the form of higher RSEs around indices of abundance leads to greater reductions in the ABC from the OFL; hence, increased survey RSE will lead to lower quotas. Results from analyses that examine the sensitivity of survey RSE to reduced trap sampling can be used to examine various funding reduction scenarios on survey RSE. Conveying this information to stakeholders and policymakers directly translates the costs of reducing survey resources in a time of funding limitations.

Another means to deal with limited funding would be to use some approach for *a priori* determining the most cost-effective sampling routes (traveling salesman problem) for which algorithms exist.

In the near future, we agree with the MARMAP/SEAMAP/SEFIS plan to eliminate the short longline collections for 2012 because we feel that the integrity of the survey will not be compromised by going to our recommended biennial sampling strategy. Reduced processing of the life history samples in the short term as long as collected and stored will not seriously compromise the survey information.

We also recommend shifting the long long-line monitoring program out of MARMAP/SEAMAP and into some type of industry/science cooperative project for reasons described above. Lastly, biological samples that will go missing from long- and short bottom-long line monitoring programs might be obtained through research set aside or CRP type projects.

Recommended Studies in Order of Priority

1. Randomization study recommendation:
    a. A CPUE standardization model needs to be done, accounting for correlation structures. A GLMM model with a random (for the cluster of 6) and a repeated effect (for the spatial surface modeling) would account for pseudo-replication due to high correlation of sites within an event and between events in close proximity to each other.
    b. Using this model for CPUE, a sensitivity analysis of the RSE versus sample size can be performed.
    c. Using this model for CPUE, estimate RSE versus sample spacing to determine the appropriate spatial allocation. Spatial thinning could be examined to determine how to reallocate samples.
    d. Based on a desired RSE of 0.2 (SAFIMP workshop 2009, page 34), if over-sampling occurs then the samples that represent 'duplicate' information could then be reallocated to priority recon sites as identified in recommendation (2). We recommend that changes proceed with caution however. For common species the current design may likely take too many samples whereas for more rare species, it may be difficult to reduce sample size and maintain RSE.
2. Explore the potential bias that adding the recon sites over time might have on CPUE indices.
    a. Annual stratum level means should be compared rather than overall means.
    b. Using the model listed in part 1a, an additional factor of recon or random event should be included and tested for significance.
3. Identify priority recon sites: Using gridded cells over the sampling universe, develop a map of the categorical probability of having hard bottom. A comparison of that with a map of the grid cells which are inadequately represented in the sampling frame would provide a ranking of priority areas for further sampling.
4. Analysis of spatial and temporal correlation structures in the data. This is useful for the CPUE standardization model and for spacing sampling sites.
5. Trap mensuration studies:
    a. Does saturation exist? Model the catch per soak time for a number of species. If saturation exists, then consider either incorporating that directly into a statistical model or alter the soak times in the field.
    b. Do competitive interactions exist? Use of the video data may provide some insight.
    c. Length-based selectivity.
6. Video mensuration studies:
    a. Examine potential for saturation of the frame in particular when reading distance is fixed.
    b. Examine the effects of turbidity and light level on the recorded counts.

References

Bacheler, N. M., Schobernd, C. M., Schobernd, Z. H., Mitchell, W. A., Berrane, D., J., Kellison, T., Reichert, M. J., M. In review. Comparison of trap and underwater video gears for indexing reef fish presence and abundance in the southeast United States.

Cochran, W. G. 1977. *Sampling Techniques*, 3rd ed. Wiley & Sons, New York.

Cameron, A.C. and P.K. Trivedi (1998). *Regression analysis of count data,* Cambridge University Press.

Dunn, D. and P. Halpin. 2009. Rugosity-based regional modeling of hard-bottom habitat. Marine Ecology Progress Series. 377: 1–11.

Millar, R. B. 1992. Estimating the size-selectivity of fishing gear by conditioning on the total catch. J. Amer. Stat. Assoc. 87: 962-968.

NOAA 2003. NOAA Fisheries Protocols: fixed gear trap, pot and camera pod surveys. Prepared by NOAA Fisheries Science Centers. December 22, 2003.

Royle, J. A. 2004. N-mixture models for estimating population size from spatially replicated counts. Biometrics, 60, 108-116.

Wileman, D. A., R. S. T. Ferro, R. Fonteyne, and R. B. Millar. 1996. Manual of methods of measuring the selectivity of towed fishing gears. ICES (International Council for the Exploration of the Sea) Cooperative Research Report 125.